

# The Multidimensional Moment Problem with Complexity Constraint

Johan Karlsson, Anders Lindquist and Axel Ringh

**Abstract.** A long series of previous papers have been devoted to the (one-dimensional) moment problem with nonnegative rational measure. The rationality assumption is a complexity constraint motivated by applications where a parameterization of the solution set in terms of a bounded finite number of parameters is required. In this paper we provide a complete solution of the multidimensional moment problem with a complexity constraint also allowing for solutions that require a singular measure added to the rational, absolutely continuous one. Such solutions occur on the boundary of a certain convex cone of solutions. In this paper we provide complete parameterizations of all such solutions. We also provide errata for a previous paper in this journal coauthored by one of the authors of the present paper.

**Mathematics Subject Classification (2010).** Primary 30E05; Secondary 42A70, 44A60, 47A57, 93A30.

**Keywords.** Moment problems, multidimensional moment problems, complexity constraints, optimization, smooth parameterization.

## 1. Introduction

The multidimensional moment problem considered in this paper amounts to finding a nonnegative measure  $d\mu$  on a compact subset  $K$  of  $\mathbb{R}^d$  solving the equations

$$c_k = \int_K \alpha_k d\mu, \quad k = 1, 2, \dots, n, \quad (1.1)$$

where  $c_1, c_2, \dots, c_n$  are given numbers and  $\alpha_1, \alpha_2, \dots, \alpha_n$  are given linearly independent basis functions defined on  $K$ . More precisely we are interested

---

This work was supported by the Swedish Foundation of Strategic Research (SSF), the Swedish Research Foundation (VR) and the Center for Industrial and Applied Mathematics (CIAM).

in measures of the type

$$d\mu(x) = \frac{P(x)}{Q(x)}dx + d\hat{\mu}(x), \quad (1.2)$$

where  $P$  and  $Q$  are nonnegative functions on  $K$  formed by linear combinations of the basis functions and  $d\hat{\mu}$  is a singular measure.

Such constraints are nonclassical and motivated by applications. An important special case is to find an absolutely continuous measure

$$d\mu(x) = \frac{P(x)}{Q(x)}dx \quad (1.3)$$

satisfying (1.1), where  $P$  and  $Q$  are as in (1.2). Clearly (1.3) is a complexity constraint depending on a finite number of parameters.

In the one-dimensional case, (generalized) moment problems with the complexity constraint (1.3) has been considered in a long series of papers [5–8, 15–17]. This study started with the *rational covariance extension problem*, which is a trigonometric moment problem with rational positive measure formulated by Kalman [22], i.e.,  $\alpha_1, \alpha_2, \dots, \alpha_n$  are the trigonometric monomials,  $e^{ikx}$ ,  $k = 0, 1, \dots, n-1$ . In this case, given moments  $c_1, c_2, \dots, c_n$  that admit solutions of (1.1), it was shown in [12, 13] that there exists a solution (1.3) for each choice of positive  $P$ , and in [9] it was established that this parameterization of the solution set is complete and smooth, i.e., the map from  $P$  to  $Q$  is a diffeomorphism. A constructive proof based on a certain family of convex optimization problems was given in [1, 3, 4]. These results were modified in steps to the case that  $\alpha_1, \alpha_2, \dots, \alpha_n$  are Herglotz kernels in [2, 14, 15, 35], leading to Nevanlinna-Pick interpolation with positive rational measure and more general moment problems with complexity constraints [5–8, 11, 18, 36].

In this paper we begin by generalizing certain results in [6] concerning moment problems over the general class of measures (1.3) to the multidimensional case ( $d > 1$ ), but we shall also take a fresh look at the case  $d = 1$ . However, when allowing  $P$  and  $Q$  to have zeros in  $K$ , the class of measures have to be extended to (1.2). Unfortunately, a key result for this case in [6] is incorrect, and we take the opportunity to provide a correction in this paper.

The multidimensional moment problem is important in many applications, such as imaging, radar, sonar, and medical diagnostics [10, 21, 32]. A series of papers by Lang and McClellan [28, 29, 31] are of special interest to us, since in a certain sense they provide an interesting overlap with the theory described above. For this reason we shall have reasons to return to some of their results in the rest of this paper.

The outline of the paper is as follows. In Section 2 we define the general multidimensional moment problem and introduce a set of dual cones that will be fundamental in the subsequent results. We present a generalization to the multidimensional case of a theorem by Krein and Nudelman [26, p. 58] on the existence of solutions to the moment problem. In Section 3 we generalize some basic results in [6] for moment problems with rational measure to the multi-variable case, and in Section 4 we introduce the basic optimization problem,

generalized to the multivariable case, and prove the basic parameterization result in the rational case. In Section 5 we extend the parameterization to the general case when, e.g.,  $P$  and  $Q$  are allowed to have zeros in  $K$ , resulting in solutions of the form (1.2). In Section 6 we consider the case when the moments are placed on the boundary of the feasible set. Section 7 is an appendix to which we have deferred some supporting results and proofs for better readability. Finally, Section 8 contains errata for [6].

## 2. The general multidimensional moment problem

Let  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$  be a set of real-valued, continuous functions defined on the compact set  $K \subset \mathbb{R}^d$ . We assume that  $K$  has an interior of dimension  $d$ , the closure of which is precisely  $K$ . The functions  $\alpha_1, \alpha_2, \dots, \alpha_n$  are assumed to be linearly independent, an assumption that will be retained throughout the paper. This condition holds in most interesting moment problems, as, for example, the trigonometric moment problem, Nevanlinna-Pick interpolation and the power moment problem.

Given real numbers  $c_1, c_2, \dots, c_n$ , we shall consider the (truncated) moment problem to determine a bounded nonnegative measure  $d\mu$  such that (1.1) holds. Whenever convenient, we shall write (1.1) in the vector form

$$c = \int_K \alpha d\mu, \quad (2.1)$$

where  $c = (c_1, c_2, \dots, c_n)'$  and  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)'$  are column vectors and prime  $(')$  denotes transpose.

The assumption that the basis functions are real can be done without loss of generality. In fact, in the case of complex basis functions (as, for example, in the trigonometric moment problem) and complex moments, we merely exchange a complex moment equation with two real moment conditions [6].

Next we define the open convex cone  $\mathfrak{P}_+ \subset \mathbb{R}^n$  of sequences  $p = (p_1, p_2, \dots, p_n)$  such that the corresponding generalized polynomial

$$P(x) = \sum_{k=1}^n p_k \alpha_k(x) \quad (2.2)$$

is positive for all  $x = (x_1, \dots, x_d) \in K$ . Moreover, we denote by  $\bar{\mathfrak{P}}_+$  its closure and by  $\partial\mathfrak{P}_+$  its boundary  $\bar{\mathfrak{P}}_+ \setminus \mathfrak{P}_+$ . It is easy to see that  $P \equiv 0$  if and only if  $p = 0$ , since  $\alpha_1, \alpha_2, \dots, \alpha_n$  are linearly independent.

Throughout the paper we assume that the zero locus of  $P$  has measure zero for each  $p \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ . Again many important moment problems, even in the multidimensional case, have this property.

Moreover, we define the dual cone

$$\mathfrak{C}_+ = \{c \in \mathbb{R}^n \mid \langle c, p \rangle > 0 \text{ for all } p \in \bar{\mathfrak{P}}_+ \setminus \{0\}\}, \quad (2.3)$$

where  $\langle c, p \rangle$  is the inner product

$$\langle c, p \rangle = \sum_{k=1}^n c_k p_k.$$

**Proposition 2.1.** *The dual cone  $\mathfrak{C}_+$  is nonempty if  $\bar{\mathfrak{P}}_+ \neq \{0\}$ .*

*Proof.* Take  $c = \int_K \alpha dx$ . Then, for any  $p \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ , we have

$$\langle c, p \rangle = \sum_{k=1}^n \int_K \alpha_k p_k dx = \int_K P dx \geq 0. \quad (2.4)$$

However, due to the continuity and linear independence,  $P \not\equiv 0$ , and hence there is always a small neighborhood in which  $P > 0$ . Consequently, the inequality in (2.4) is strict for all  $p \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ , and hence  $c \in \mathfrak{C}_+$ .  $\square$

The dual cone  $\mathfrak{C}_+$  is also an open cone, and we denote by  $\bar{\mathfrak{C}}_+$  its closure and by  $\partial\mathfrak{C}_+$  its boundary.

**Proposition 2.2.** *Any  $c \in \mathbb{R}^n$  satisfying (2.1) for some nonnegative measure  $d\mu$  belongs to  $\bar{\mathfrak{C}}_+$ .*

*Proof.* If  $c$  satisfies (2.1), then

$$\langle c, p \rangle = \sum_{k=1}^n p_k \int_K \alpha_k d\mu = \int_K P d\mu \geq 0$$

for all  $p \in \bar{\mathfrak{P}}_+$ . Hence  $c \in \bar{\mathfrak{C}}_+$ , as claimed.  $\square$

The converse also holds. The proof of the following generalization to the multidimensional case of a result in Krein and Nudelman [26, p. 58] is based on Theorem 5.1 and will be deferred to Section 6.

**Theorem 2.3.** *Suppose that  $\mathfrak{P}_+ \neq \emptyset$ . Then for any  $c \in \bar{\mathfrak{C}}_+$  there exists a bounded nonnegative measure  $d\mu$  such that (2.1) holds.*

This theorem ensures that the space

$$\mathfrak{M}_c = \left\{ d\mu \geq 0 \mid \int_K \alpha d\mu = c \right\} \quad (2.5)$$

of bounded measures is nonempty for all  $c \in \bar{\mathfrak{C}}_+$ . However, in general, there are infinitely many solutions, and the extreme points of  $\mathfrak{M}_c$  are of particular interest.

**Proposition 2.4.** *Suppose that  $\mathfrak{P}_+ \neq \emptyset$ . Then, for all  $c \in \bar{\mathfrak{C}}_+$  there is a  $d\mu \in \mathfrak{M}_c$  that is a discrete measure with support in at most  $n$  points in  $K$ .*

*Proof.* Following the proof of Proposition 6 in [24], we use the Krein-Millman Theorem [19, 25] to show the existence of an extreme point, and then we show that any extreme point has the claimed properties. The space  $\mathfrak{M}_c$  is the intersection of a positive cone and a closed subspace, and hence it is convex and closed. Let  $p \in \mathfrak{P}_+$ . Then  $P(x) \geq \varepsilon$  for some  $\varepsilon > 0$ , and hence  $\langle c, p \rangle = \int_K P d\mu \geq \varepsilon \mu(K)$ . Therefore the norm (total variation) of the elements of

$\mathfrak{M}_c$  is bounded by  $\langle c, p \rangle / \varepsilon$ , which implies that  $\mathfrak{M}_c$  is compact in the weak\* topology [19, p. 19]. Then, since  $\mathfrak{M}_c$  is a compact convex set in a locally convex topological linear space, it is the closure of the convex hull of its extreme points [19, 25]. Since the set  $\mathfrak{M}_c$  is nonempty (Theorem 2.3), it has at least one extreme point.

We want to prove that the extreme points of  $\mathfrak{M}_c$  have support in at most  $n$  points. To this end, suppose the contrary. Then there is an extreme point  $d\mu = \sum_{k=1}^{n+1} \beta_k d\mu_k$  for which the measures  $d\mu_k \geq 0$  have distinct support and  $\beta_k > 0$ ,  $k = 1, \dots, n+1$ , and therefore

$$c = \int_K \alpha d\mu = \int_K \alpha [d\mu_1 \cdots d\mu_{n+1}] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n+1} \end{bmatrix}.$$

Since  $\int_K \alpha [d\mu_1 \cdots d\mu_{n+1}] \in \mathbb{R}^{n \times (n+1)}$  has linearly dependent columns, there is a nontrivial affine set of solutions  $\{\beta_k\}$ , which contradicts the assumption that  $d\mu$  is an extreme point of  $\mathfrak{M}_c$ . Hence the extreme points of  $\mathfrak{M}_c$  has support in at most  $n$  points.  $\square$

**Corollary 2.5.** *The extreme points of  $\mathfrak{M}_c$  have support in at most  $n$  points.*

In several of the most important one-dimensional moment problems there are matrix tests available to check that  $c \in \mathfrak{C}_+$ . For example, in the trigonometric moment problem, we need to check that the corresponding Toeplitz matrix is positive definite, and, in the Nevanlinna-Pick interpolation, that the Pick matrix is positive definite. There is also a matrix test for the power moment problem in terms of a Hankel matrix [26].

In the multidimensional case, checking that  $c \in \mathfrak{C}_+$  is more complicated, but the following result might provide some help.

**Proposition 2.6.** *Let  $c_0 \in \mathfrak{C}_+$  be arbitrary.<sup>1</sup> Then the optimization problem*

$$\min_{p \in \mathfrak{P}_+} \langle c, p \rangle \quad \text{subject to } \langle c_0, p \rangle = 1 \quad (2.6)$$

*has a solution with minimal value  $V$ . Moreover, (i)  $V > 0$  if and only if  $c \in \mathfrak{C}_+$ , (ii)  $V = 0$  if and only if  $c \in \partial \mathfrak{C}_+$ , and (iii)  $V < 0$  if and only if  $c \notin \bar{\mathfrak{C}}_+$ .*

*Proof.* The constraint  $\langle c_0, p \rangle = 1$  is a hyperplane that does not pass through the vertex of the cone  $\mathfrak{P}_+$  where  $p = 0$ . Since, in addition,  $c_0 \in \mathfrak{C}_+$ , the hyperplane has a compact intersection with  $\mathfrak{P}_+$ , and therefore the linear functional  $\langle c, p \rangle$  has always a minimum there. Then the rest of the proposition follows directly from the definition (2.3).  $\square$

---

<sup>1</sup>For example  $c_0 = \int_K \alpha dx \in \mathfrak{C}_+$  as in the proof of Proposition 2.1.

As a preamble to the solution of the fundamental optimization problem of Section 5, it is instructive to consider the dual problem of (2.6). Differentiating the Lagrangian

$$L(p; d\mu, \lambda) = \langle c, p \rangle - \int_K P d\mu - \lambda(\langle c_0, p \rangle - 1),$$

where the nonnegative bounded measure<sup>2</sup>  $d\mu \in C(K)^*$  and  $\lambda$  are Lagrange multipliers, we obtain

$$c - \hat{c} - \lambda c_0 = 0 \quad \text{where } \hat{c} := \int_K \alpha d\mu, \quad (2.7)$$

and the complementary slackness condition

$$\langle \hat{c}, p \rangle = \int_K P d\mu = 0, \quad (2.8)$$

which implies that  $\hat{c} \in \partial\mathfrak{C}_+$ . Since  $P \geq 0$  on  $K$  and  $d\mu$  is a nonnegative measure, (2.8) also implies that  $d\mu$  is either identically zero or a singular measure with support in the set of zeros of  $P$ . Inserting the stationarity condition (2.7) in the Lagrangian we get the dual functional [30]

$$\varphi(d\mu, \lambda) = \langle \hat{c}, p \rangle + \lambda \langle c_0, p \rangle - \int_K P d\mu - \lambda \langle c_0, p \rangle + \lambda = \lambda.$$

Consequently, the dual problem is

$$\max_{\lambda \in \mathbb{R}} \lambda \quad \text{subject to } c - \lambda c_0 \in \partial\mathfrak{C}_+ \quad (2.9)$$

with optimal value  $\hat{\lambda}$ . Since these are convex optimization problems and  $\mathfrak{P}_+ \neq \emptyset$ , there is no duality gap [30], and therefore  $\hat{\lambda} = V$ . From this we can construct an alternative proof of Proposition 2.6. In fact, since  $\hat{c} \in \partial\mathfrak{C}_+$ , it follows from (2.7) that  $c \in \partial\mathfrak{C}_+$  is equivalent to  $\lambda c_0 \in \partial\mathfrak{C}_+$ , which in turn holds if and only if  $\lambda = 0$ , since  $c_0 \in \mathfrak{C}_+$ . Increasing  $\lambda$  through positive values brings  $c = \hat{c} + \lambda c_0$  into  $\mathfrak{C}_+$ , whereas negative values of  $\lambda$  brings  $c = \hat{c} + \lambda c_0$  outside  $\mathfrak{C}_+$ . This could be compared with the test procedure suggested in [27, p. 3.2.4].

Using a homotopy approach, it was shown in [16] that a certain differential equation has an exponentially attractive point if and only if  $c \in \mathfrak{C}_+$ . Such a procedure might be preferable from a computational point of view.

### 3. Solutions with rational positive measure

We begin by considering *rational positive measures* of the type

$$d\mu = \frac{P}{Q} dx, \quad p, q \in \mathfrak{P}_+ \quad (3.1)$$

and define the moment map

$$f^p(q) = \int_K \alpha \frac{P}{Q} dx. \quad (3.2)$$

---

<sup>2</sup>The dual space of  $C(K)$ , denoted by  $C(K)^*$ , is the space of bounded signed measures [30].

The following condition is instrumental in ensuring the existence of an interpolating measure of the form (3.1) for each given  $p \in \mathfrak{P}_+$ , which, as we will see, leads to a complete and smooth parameterization in terms of rational measures.

**Condition 3.1.** *The cone  $\mathfrak{P}_+$  is nonempty and has the property*

$$\int_K \frac{1}{Q} dx = \infty \quad \text{for all } q \in \partial\mathfrak{P}_+. \quad (3.3)$$

The following modification of Theorem 2.3 is a simple corollary of Theorem 3.4 below, but we state it already here to motivate our interest in rational measures (3.1).

**Proposition 3.2.** *Suppose that Condition 3.1 holds. Then there is a  $d\mu \in \mathfrak{M}_c$  of the form (3.1) satisfying (2.1) if and only if  $c \in \mathfrak{C}_+$ .*

This is immediate by taking any  $p \in \mathfrak{P}_+$  in Corollary 3.5.

**Lemma 3.3.** *Suppose that Condition 3.1 holds and that  $p \in \mathfrak{P}_+$ . Then the map  $f^p : \mathfrak{P}_+ \rightarrow \mathfrak{C}_+$  is proper, i.e., the inverse image  $(f^p)^{-1}(C)$  is compact for every compact  $C \subset \mathfrak{C}_+$ .*

*Proof.* We first prove that  $(f^p)^{-1}(C)$  is bounded. To this end, first note that the set  $\{(c, q) \mid c \in C, q \in \mathfrak{P}_+, \|q\|_\infty = 1\}$  is compact, and hence the bilinear form  $\langle c, q \rangle$  has a minimum  $\varepsilon$  there, where  $\varepsilon > 0$  since  $C \subset \mathfrak{C}_+$ . Hence  $\langle c, q \rangle \geq \varepsilon \|q\|_\infty$ . However,

$$\langle f^p(q), q \rangle = \int_K P dx =: \kappa$$

is constant. Consequently  $\|q\|_\infty \leq \kappa/\varepsilon$  for any  $q \in (f^p)^{-1}(C)$ , proving boundedness. If  $(f^p)^{-1}(C)$  is empty or finite, it is trivially compact, so let us assume that it contains infinitely many points. Let  $(q^k)$  be any sequence in  $\mathfrak{P}_+$  such that  $f^p(q^k) \in C$  for all  $k$ . Since  $(f^p)^{-1}(C)$  is bounded,  $(q^k)$  has a cluster point  $\hat{q}$  in the closure  $\bar{\mathfrak{P}}_+$ . Compactness of  $(f^p)^{-1}(C)$  now follows from the fact that  $\hat{q} \notin \partial\mathfrak{P}_+$ . In fact, if  $\hat{q} \in \partial\mathfrak{P}_+$  were the case, we would have

$$\langle f^p(\hat{q}), p \rangle = \int_K \frac{P^2}{Q} dx = \infty$$

by the assumptions  $p \in \mathfrak{P}_+$  and Condition 3.1, which contradicts the fact that  $\sup_k \langle f^p(q^k), p \rangle \leq \sup_{c \in C} \langle c, p \rangle$  is finite.  $\square$

**Theorem 3.4.** *Suppose that Condition 3.1 holds and that  $p \in \mathfrak{P}_+$ . Then the map  $f^p : \mathfrak{P}_+ \rightarrow \mathfrak{C}_+$  is a diffeomorphism between  $\mathfrak{P}_+$  and  $\mathfrak{C}_+$ .*

*Proof.* Since  $p \in \mathfrak{P}_+$ , the Jacobian

$$\frac{\partial f^p}{\partial q} = - \int_K \alpha \frac{P}{Q^2} \alpha' dx < 0 \quad (3.4)$$

on all of  $\mathfrak{P}_+$ . In fact, for  $a \in \mathbb{R}^n$ , the quadratic form

$$a' \frac{\partial f^p}{\partial q} a = - \int_K (a' \alpha)^2 \frac{P}{Q^2} dx = 0$$

if and only if  $a = 0$ . Since  $f^p$  is also proper (Lemma 3.3), it follows from Hadamard's global inverse function theorem [20] that the map  $f^p$  is a diffeomorphism.  $\square$

**Corollary 3.5.** *Suppose that Condition 3.1 holds. Then the moment equations*

$$c_k = \int_K \alpha_k \frac{P}{Q} dx, \quad k = 0, 1, \dots, n \quad (3.5)$$

*have a unique solution  $q \in \mathfrak{P}_+$  for each  $(c, p) \in \mathfrak{C}_+ \times \mathfrak{P}_+$ .*

*Remark 3.6.* As demonstrated in [8], Condition 3.1 holds in the one-dimensional case if the basis functions  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  are Lipschitz continuous. In fact, if  $q \in \partial\mathfrak{P}_+$ , there is an  $x_0 \in K$  such that  $Q(x_0) = 0$ . Therefore, if  $Q$  is Lipschitz continuous, there exists an  $\varepsilon > 0$  and a  $\kappa > 0$  such that  $Q(x) \leq \kappa|x - x_0|$  for all  $x \in [x_0 - \varepsilon, x_0 + \varepsilon]$ , and hence

$$\int_K \frac{dx}{Q} \geq \int_{x_0 - \varepsilon}^{x_0 + \varepsilon} \frac{dx}{Q} \geq \frac{1}{\kappa} \int_{x_0 - \varepsilon}^{x_0 + \varepsilon} \frac{dx}{|x - x_0|} = \infty.$$

This is a very mild condition, since any reasonable moment problem encountered in applications would have Lipschitz continuous basis functions. Also, if  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  is a Chebyshev system (or T-system)<sup>3</sup> and contains a constant function, then after a reparameterization the basis functions will be Lipschitz continuous [26, p. 37].

In the multidimensional case, the situation is a bit trickier. As was noted in [16, p. 819], Condition 3.1 always holds if  $K$  is an interval in  $\mathbb{R}^2$  and the basis functions  $(\alpha_k)$  are twice differentiable and doubly periodic. However, as the following example from [28] illustrates, this does not hold for  $d \geq 3$ .

*Example 3.7.* Let  $K := [-\pi, \pi]^d$  and  $\alpha_k(x) = \cos x_k$ ,  $k = 1, 2, \dots, d$ , and set

$$Q(x) = \sum_{k=1}^d (1 - \cos x_k),$$

which corresponds to a  $q \in \partial\mathfrak{P}_+$  since  $Q(0) = 0$ . However, this is the only zero of  $Q$ , and hence, in checking Condition 3.1, we need only consider a small neighborhood  $D_\varepsilon = \{x \in K \mid \|x\| \leq \varepsilon\}$  of  $x = 0$ . A series expansion shows that for sufficiently small  $\varepsilon > 0$  we have  $1 - \cos x_k \geq x_k^2/4$  on all of  $D_\varepsilon$ , and hence

$$\int_{D_\varepsilon} \frac{dx}{Q} \leq \int_{D_\varepsilon} \frac{dx}{\sum_{k=1}^d x_k^2/4}.$$

Changing to spherical coordinates, this becomes

$$\int_{D_\varepsilon} \frac{dx}{Q} \leq \int_{\varphi_{d-1}=0}^{2\pi} \int_{\varphi=0}^{\pi} \int_{r=0}^{\varepsilon} \frac{4}{r^2} r^{d-1} \sin \varphi \, dr \, d\varphi \, d\varphi_{d-1},$$

where  $\varphi := (\varphi_1, \dots, \varphi_{d-2})$  and  $\sin \varphi := \sin^{d-2} \varphi_1 \cdots \sin \varphi_{d-2}$ . This is clearly finite for  $d \geq 3$ , and hence Condition 3.1 does not hold in these cases.

<sup>3</sup>A set of real functions  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  on an interval  $[a, b]$  is called a Chebyshev system if any nonzero linear combination  $P(x) = \sum_{k=1}^n p_k \alpha_k(x)$  has at most  $n$  zeros [26, p. 31].



*Example 3.8.* Next consider a one-dimensional case where  $\alpha$  is not Lipschitz continuous. Let  $\alpha(x) = (1, x^{1/3}, x^{2/3})$  and  $K = [-1, 1]$ . Then  $q = (0, 0, 3) \in \partial\mathfrak{P}_+$ . In fact,  $Q(x) = 3x^{2/3}$ , and hence  $Q(0) = 0$ . However,

$$\int_{-1}^1 \frac{dx}{Q(x)} = \int_{-1}^1 \frac{1}{3} x^{-2/3} dx = 2 < \infty, \quad (3.6)$$

so Condition 3.1 is not satisfied. In this case,  $f^p : \mathfrak{P}_+ \rightarrow \mathfrak{C}_+$  is not a diffeomorphism, so we have a counterexample to the statement of Theorem 3.4 with Condition 3.1 removed. In fact,  $f^p$  is not even continuous in all points. Take  $p := (1, 0, 0) \in \mathfrak{P}_+$  and a sequence  $q_k := 3(k^{-2}, 2k^{-1}, 1) \in \mathfrak{P}_+$ , which tends to  $q_\infty := (0, 0, 3)$  as  $k \rightarrow \infty$ . Then  $P = 1$  and  $Q_k(x) = 3(x^{1/3} + k^{-1})^2$ , and hence  $P/Q_k \rightarrow P/Q_\infty = \frac{1}{3}x^{-2/3}$  for all  $x \neq 0$ . However, by (3.6),  $f_1^p(q_\infty) = \int_{-1}^1 P/Q_\infty dx = 2$ , whereas

$$f_1^p(q_k) = \int_{-1}^1 \frac{P}{Q_k} dx = \frac{1}{3} \int_{-1}^1 (x^{1/3} + k^{-1})^{-2} dx = \int_{-1}^1 \frac{x^2}{(x^2 + k^{-1})^{-2}} dx = \infty$$

for all  $k \geq 1$ . Consequently,  $f^p$  is not continuous in  $q = (0, 0, 3)$ . Moreover, we observe that the substitution  $y = x^{1/3}$  transforms the basis to  $\alpha(x) = (1, x, x^2)$ , which is Lipschitz continuous. However, now  $P(x) = x^2$ , and hence  $p \in \partial\mathfrak{P}_+$ , so this is not a counterexample to Theorem 3.4. We may even modify this example so that  $f_1^p(q_k) < \infty$  for all finite  $k$ . To this end, choose  $q_k := 3(k^{-2} + k^{-4}, 2k^{-1}, 1)$ , which again tends to  $q_\infty := (0, 0, 3)$ . However, it can now be shown that

$$f_1^p(q_k) = \int_{-1}^1 \frac{P}{Q_k} dx \rightarrow 2 + \pi > 2 = f_1^p(q_\infty)$$

as  $k \rightarrow \infty$ , which again shows that  $f^p$  is not continuous.

#### 4. The optimization problem

Next, given the moment map (3.2), following [5, 6] we construct the 1-form

$$\begin{aligned} \omega = \langle c - f^p(q), dq \rangle &= \sum_{k=1}^n c_k dq_k - \int_K \sum_{k=1}^n \alpha_k \frac{P}{Q} dq_k dx \\ &= \langle c, dq \rangle - \int_K \frac{P}{Q} dQ dx \end{aligned}$$

on  $\mathfrak{P}_+$ . Taking the exterior derivative (on  $\mathfrak{P}_+$ ) we obtain

$$d\omega = \int_K \frac{P}{Q^2} dQ \wedge dQ dx = 0,$$

i.e., the 1-form  $\omega$  is closed. Therefore, since  $\mathfrak{P}_+$  is an open convex set and hence star-shaped in any point,  $\omega$  is exact by the Poincaré Lemma [39, pp. 92-94]. This means that there exists a smooth function  $\mathbb{J}_p^c$  on  $\mathfrak{P}_+$  such that

$$\mathbb{J}_p^c(q_1) - \mathbb{J}_p^c(q_0) = \int_{q_0}^{q_1} \omega = \int_{q_0}^{q_1} \langle c, dq \rangle - \int_{q_0}^{q_1} \int_K \frac{P}{Q} dQ dx,$$

with the integral being independent of the path between two endpoints. Computing the path integral, one finds that

$$\mathbb{J}_p^c(q) = \langle c, q \rangle - \int_K P \log Q dx \quad (4.1)$$

modulo a constant of integration. Then, for each  $(c, p) \in \mathfrak{C}_+ \times (\bar{\mathfrak{P}}_+ \setminus \{0\})$  we extend the functional (4.1) to a map  $\bar{\mathfrak{P}}_+ \rightarrow \mathbb{R} \cup \{\infty\}$ .

**Lemma 4.1.** *Let  $(c, p) \in \mathfrak{C}_+ \times (\bar{\mathfrak{P}}_+ \setminus \{0\})$ . Then the functional  $\mathbb{J}_p^c : \bar{\mathfrak{P}}_+ \rightarrow \mathbb{R}$  is strictly convex.*

*Proof.* First observe that

$$\frac{\partial \mathbb{J}_p^c}{\partial q} = c - f^p(q),$$

and consequently, in view of (3.4), the Hessian  $\mathbb{H}(q)$  of  $\mathbb{J}_p^c$  is given by

$$\mathbb{H}(q) = -\frac{\partial f^p}{\partial q}(q) = \int_K \alpha \frac{P}{Q^2} \alpha' dx,$$

where the integrand  $P/Q^2$  is nonnegative. Since by assumption the zero locus of  $P$  has measure zero,  $P/Q^2$  is zero at most on a subset of  $K$  of measure zero, and consequently  $\mathbb{H}(q) > 0$  for the same reason as in the proof of Theorem 3.4. This implies that  $\mathbb{J}_p^c$  is strictly convex.  $\square$

**Lemma 4.2.** *The map  $\mathbb{J}_p^c : \bar{\mathfrak{P}}_+ \rightarrow \mathbb{R}$  is lower semi-continuous on all of  $\bar{\mathfrak{P}}_+$ .*

*Proof.* Let  $(q_k)$  be a sequence in  $\bar{\mathfrak{P}}_+$  converging to  $q \in \bar{\mathfrak{P}}_+$  in  $L_\infty$ -norm. Since the functions  $(Q_k)$  and  $Q$  are continuous on a compact set  $K$ , the convergence  $Q_k \rightarrow Q$  as  $k \rightarrow \infty$  is uniform. Since  $Q$  is a continuous function on a compact set,  $\max_x Q < \infty$ . Moreover, since the convergence is uniform, we have that  $\sup_k \max_x Q_k < \infty$ . Hence there is an  $M$  such that  $\max_x Q \leq M$  and  $\sup_k \max_x Q_k \leq M$ , and thus

$$-\log\left(\frac{Q}{M}\right) \geq 0 \quad \text{and} \quad -\log\left(\frac{Q_k}{M}\right) \geq 0, \quad k = 1, 2, \dots$$

Therefore, by Fatou's lemma,

$$-\int_K \log\left(\frac{Q}{M}\right) dx \leq \liminf_{k \rightarrow \infty} -\int_K \log\left(\frac{Q_k}{M}\right) dx$$

since  $Q_k \rightarrow Q$  pointwise. Consequently,  $\mathbb{J}_p^c(q) \leq \liminf_{k \rightarrow \infty} \mathbb{J}_p^c(q_k)$ , proving that  $\mathbb{J}_p^c$  is lower semicontinuous.  $\square$

**Lemma 4.3.** *The sublevel sets of  $\mathbb{J}_p^c$  are compact, i.e., the inverse image  $(\mathbb{J}_p^c)^{-1}(\infty, r]$  is compact for all values of  $r$ .*

*Proof.* By Lemma 7.1 in the appendix we have

$$r \geq \mathbb{J}_p^c(q) \geq \epsilon_c \|Q\|_\infty - \epsilon_p \log \|Q\|_\infty,$$

and by comparing linear and logarithmic growth we see that the sublevel sets are bounded from both above and below. Since they are sublevel sets of a

lower semi-continuous function (Lemma 4.2), they are closed. Hence they are compact.  $\square$

**Theorem 4.4.** *Let  $(c, p) \in \mathfrak{C}_+ \times (\bar{\mathfrak{P}}_+ \setminus \{0\})$ . Then the functional (4.1) has a unique minimum in  $\bar{\mathfrak{P}}_+$ . Moreover, the map  $g^c : \bar{\mathfrak{P}}_+ \setminus \{0\} \rightarrow \bar{\mathfrak{P}}_+$  sending  $p$  to the corresponding minimizer  $\hat{q}$  is continuous and injective. If, in addition, Condition 3.1 holds and  $p \in \mathfrak{P}_+$ , then the minimizer  $\hat{q}$  belongs to  $\mathfrak{P}_+$ .*

*Proof.* By Lemma 4.3 there exists a minimizer  $\hat{q} \in \bar{\mathfrak{P}}_+$ . Since  $\mathbb{J}_p^c$  is strictly convex (Lemma 4.1), this minimizer is unique. Injectivity of the function  $g^c$  follows from a trivially modified version of the proof of Lemma 2.4 in [6]. Continuity of  $g^c$  is proved in the appendix (Proposition 7.3). If Condition 3.1 holds and  $p \in \mathfrak{P}_+$ , the functional (4.1) has a stationary point in the open cone  $\mathfrak{P}_+$  (Theorem 3.4), which must then be identical to  $\hat{q}$ .  $\square$

We note in passing that, for  $(c, p) \in \mathfrak{C}_+ \times \mathfrak{P}_+$ , the convex optimization problem to minimize  $\mathbb{J}_p^c$  over all  $q \in \mathfrak{P}_+$  is the dual of the problem to maximize

$$\mathbb{I}_p(\Phi) = \int_K P(x) \log \Phi(x) dx \quad (4.2)$$

over all  $\Phi \in \mathcal{F}_+$  satisfying the moment condition

$$\int_K \alpha(x) \Phi(x) dx = c, \quad (4.3)$$

where  $\mathcal{F}_+$  is the class of positive functions in  $L_1(K)$ . In fact, we have the following duality result.

**Theorem 4.5.** *Suppose that  $(c, p) \in \mathfrak{C}_+ \times \mathfrak{P}_+$  and that Condition 3.1 holds. Then the optimization problem to maximize (4.2) over all  $\Phi \in \mathcal{F}_+$  satisfying the moment condition (4.3) has a unique solution*

$$\hat{\Phi} = \frac{P}{\hat{Q}},$$

where  $\hat{q}$  is the unique minimizer of  $\mathbb{J}_p^c$ . Moreover,

$$\mathbb{I}_p(\hat{\Phi}) = \mathbb{J}_p^c(\hat{q}) + \int_K P(\log P - 1) dx.$$

*Proof.* Given Lagrange multipliers  $q = (q_1, q_2, \dots, q_n)$ , form the Lagrangian

$$\begin{aligned} L(\Phi, q) &= \mathbb{I}_p(\Phi) + \sum_{k=1}^n q_k \left( c_k - \int_K \alpha_k \Phi dx \right) \\ &= \mathbb{I}_p(\Phi) + \langle c, q \rangle - \int_K Q \Phi dx, \end{aligned}$$

which is finite for any fixed  $q \in \mathfrak{P}_+$ . Setting the Fréchet differential

$$\delta L(\Phi, q; \delta \Phi) = \int_K \left( \frac{P}{\Phi} - Q \right) \delta \Phi dx = 0$$

for all  $\delta\Phi$ , we obtain the stationary point  $\Phi = P/Q$ , which inserted into the Lagrangian yields

$$L(P/Q, q) = \mathbb{J}_p^c(q) + \int_K P(\log P - 1)dx.$$

Since  $\Phi \mapsto L(\Phi, q)$  is concave for any  $q \in \mathfrak{P}_+$ , we have  $L(\Phi, q) \leq L(P/Q, q)$  for all  $\Phi \in \mathcal{F}_+$ . However, by Theorem 4.4, there is a unique  $q \in \mathfrak{P}_+$  such that  $\Phi := P/Q$  satisfies (4.3), namely  $\hat{q}$ , and hence

$$\mathbb{I}_p(\Phi) = L(\Phi, \hat{q}) \leq L(P/\hat{Q}, \hat{q}) = \mathbb{I}_p(\hat{\Phi})$$

for all  $\Phi \in \mathcal{F}_+$  satisfying (4.3), proving the required optimality.  $\square$

The choice  $P \equiv 1$  yields the *maximum entropy* solution of the moment problem. This duality has been extensively discussed in [1–5, 7, 8, 18] in the one-dimensional case. In the special case of trigonometric basis functions, the multidimensional case was already covered in [28, 29, 31], but in a more general framework of weighted maximum entropy optimization which does not consider parameterization of rational solutions and related issues, something that is important in our present context.

The duality result of Theorem 4.5 relies on Condition 3.1 and will be reformulated in Section 5, where this condition will not be required. In this case the optimal solution may contain a singular part, that is, be of the form (1.2) (Theorem 5.2).

**Theorem 4.6.** *Suppose that Condition 3.1 holds and  $c \in \mathfrak{C}_+$ . Then the map  $g^c : \mathfrak{P}_+ \rightarrow \mathfrak{P}_+$ , restricted to  $\mathfrak{P}_+$ , is a diffeomorphism onto its image  $\mathfrak{Q}_+$ .*

*Proof.* Since  $g^c$  is continuous and injective (Theorem 4.4),  $\mathfrak{Q}_+$  is an open set of the same dimension as  $\mathfrak{P}_+$ . By definition,  $g^c : \mathfrak{P}_+ \rightarrow \mathfrak{Q}_+$  is also surjective. Next define the function

$$\varphi(p, q) = c - \int_K \alpha \frac{P}{Q} dx.$$

Then the moment equations (stationarity condition) can be written  $\varphi(p, q) = 0$ . Since

$$\frac{\partial \varphi}{\partial q} = \int_K \alpha \frac{P}{Q^2} \alpha' dx$$

is positive definite on all of  $\mathfrak{P}_+ \times \mathfrak{P}_+$ , the Implicit Function Theorem implies that  $q = g^c(p)$  where  $g^c$  is continuously differentiable. Moreover,

$$\frac{\partial g^c}{\partial p}(p) = - \left[ \int_K \alpha \frac{P}{Q^2} \alpha' dx \right]^{-1} \frac{\partial \varphi}{\partial p}(p, g^c(p))$$

is positive definite, since

$$\frac{\partial \varphi}{\partial p}(p, q) = - \int_K \alpha \frac{1}{Q} \alpha' dx$$

is negative definite. Hence, by the Inverse Function Theorem, the inverse function  $(g^c)^{-1}$  is also continuously differentiable. Consequently,  $g^c : \mathfrak{P}_+ \rightarrow \mathfrak{Q}_+$  is a diffeomorphism.  $\square$

Together with Corollary 3.5, Theorem 4.6 yields a complete parameterization of all solutions of the rational moment equations (3.5). This generalizes to the multidimensional case the corresponding results in [6], which in turn are generalizations of the results in [9].

If  $q \in \partial\mathfrak{P}_+$ , it follows that  $q \notin \mathfrak{Q}_+ := g^c(\mathfrak{P}_+) \subset \mathfrak{P}_+$ , and hence, by Theorem 4.6,  $p \notin \mathfrak{P}_+$ , and thus  $p \in \partial\mathfrak{P}_+$ . This yields the following corollary.

**Corollary 4.7.** *Suppose that Condition 3.1 holds,  $c \in \mathfrak{C}_+$ , and  $P/Q$  satisfies the rational moment condition (3.5). Then  $q \in \partial\mathfrak{P}_+$  implies that  $p \in \partial\mathfrak{P}_+$ .*

Note, however, that the converse is not true. In fact, a  $q \in \partial\mathfrak{Q}_+$  could be contained in  $\mathfrak{P}_+$  as the following simple one-dimensional example shows. Let  $K = [-\pi, \pi]$ ,  $\alpha_1 = 1$  and  $\alpha_2 = \cos x$ , and suppose that  $P(x) = 1 - \cos x$  and  $Q(x) \equiv 1$ . Then  $c = (1, -1/2)$ . Since  $c$  has a positive definite Toeplitz matrix,  $c \in \mathfrak{C}_+$ . Moreover, since  $P(0) = 0$ ,  $p \in \partial\mathfrak{P}_+$ , and therefore, by Theorem 4.6,  $q = (1, 0) \in \partial\mathfrak{Q}_+$ . However, clearly  $q \in \mathfrak{P}_+$ .

Moreover, again Condition 3.1 is crucial. In fact, in Example 3.8, where this condition does not hold,  $q \in \partial\mathfrak{P}_+$  whereas  $p \in \mathfrak{P}_+$  is in the interior. However, under the variable substitution  $y = x^{1/3}$ , which makes the basis Lipschitz continuous so that Condition 3.1 holds,  $p$  moves to the boundary  $\partial\mathfrak{P}_+$ .

## 5. Solutions on the boundary

If Condition 3.1 holds and  $p \in \mathfrak{P}_+$ , then  $\mathbb{J}_p^c$  has a unique minimum in the open cone  $\mathfrak{P}_+$ , which solves the moment equations (3.5). On the other hand, if these conditions are not satisfied, the minimizer may end up on the boundary  $\partial\mathfrak{P}_+$ , leading to complications described in [33, 34] for the special case of rational covariance extension. Therefore, in the present more general situation, the constraint  $Q(x) \geq 0$  becomes essential for solving the optimization problem.

**Theorem 5.1.** *Let  $(c, p) \in \mathfrak{C}_+ \times (\tilde{\mathfrak{P}}_+ \setminus \{0\})$ . Then there exists a unique pair  $(\hat{c}, \hat{q}) \in \partial\mathfrak{C}_+ \times (\tilde{\mathfrak{P}}_+ \setminus \{0\})$  such that*

$$c_k = \int_K \alpha_k \frac{P}{\hat{Q}} dx + \hat{c}_k, \quad k = 0, 1, \dots, n. \quad (5.1a)$$

Here

$$\hat{c}_k = \int_K \alpha_k d\hat{\mu}, \quad k = 0, 1, \dots, n, \quad (5.1b)$$

where  $d\hat{\mu}$  is a (not necessarily unique) singular bounded nonnegative measure such that  $\text{supp}(d\hat{\mu}) \subset \text{null}(\hat{Q})$ , i.e., the support of the discrete measure  $d\hat{\mu}$  is contained in the set of zeros of  $\hat{Q}$ . The vector  $\hat{q}$  is the unique minimizer of (4.1) over  $\tilde{\mathfrak{P}}_+$ .

*Proof.* Since  $Q \in C(K)$ , Lagrange relaxation leads to the Lagrangian

$$L(q, \mu) = \mathbb{J}_p^c(q) - \int_K Q d\mu, \quad (5.2)$$

where  $q \in \bar{\mathfrak{P}}_+$ , and where the Lagrange multiplier  $d\mu \in C(K)^*$  is a nonnegative measure. Then setting

$$\frac{\partial L}{\partial q_k}(q, \mu) = c_k - \int_K \alpha_k \frac{P}{Q} dx - \int_K \alpha_k d\mu, \quad k = 1, 2, \dots, n,$$

equal to zero, we obtain (5.1) for the saddle point  $(\hat{q}, \hat{\mu})$  [30]. By Theorem 4.4, the minimizer  $\hat{q}$  is unique. Then it is seen from (5.1a) that  $\hat{c}$  is also unique. Now, by complementary slackness [30, Theorem 1, p. 217],

$$\int_K \hat{Q} d\hat{\mu} = 0. \quad (5.3)$$

Since the optimization problem is convex with a strictly feasible point, the KKT conditions (i)  $\hat{q} \in \bar{\mathfrak{P}}_+$ , (ii)  $d\hat{\mu}$  nonnegative bounded measure in  $C(K)^*$ , (iii) (5.1) and (iv) (5.3) are necessary and sufficient for optimality. Since  $\hat{Q} \geq 0$  and  $d\hat{\mu} \geq 0$  on  $K$ , (5.3) can only hold if the support of  $d\hat{\mu}$  is contained in the set of zeros of  $\hat{Q}$ . However, such zeros exist only if  $\hat{q} \in \partial\bar{\mathfrak{P}}_+$ .

It remains to show that  $\hat{c} \in \partial\mathfrak{C}_+$ . To this end, first note that, in view of the representation (5.1b),  $\hat{c} \in \bar{\mathfrak{C}}_+$  (Proposition 2.2). Moreover, by (5.3),

$$\langle \hat{c}, \hat{q} \rangle = \int_K \hat{Q} d\hat{\mu} = 0,$$

and hence  $\hat{c} \in \partial\mathfrak{C}_+$ . □

We are now in a position to generalize the duality relation from Theorem 4.5 to the more general setting where Condition 3.1 is no longer required. Thus we extend the domain of the objective function  $\mathbb{I}_p$  of the primal problem to include any nonnegative measure  $d\mu$  and define

$$\mathbb{I}_p(d\mu) = \int_K P(x) \log \Phi(x) dx, \quad (5.4)$$

where  $d\mu = \Phi(x)dx + d\hat{\mu}$  is the unique Lebesgue decomposition of the measure [38].

**Theorem 5.2.** *Let  $(c, p) \in \mathfrak{C}_+ \times (\bar{\mathfrak{P}}_+ \setminus \{0\})$  be given. Then the problem to maximize (5.4) over the set of nonnegative measures of bounded variation, subject to the moment condition (1.1), has a solution on the form*

$$d\mu = \frac{P(x)}{\hat{Q}(x)} dx + d\hat{\mu},$$

where  $\hat{q} \in \bar{\mathfrak{P}}_+ \setminus \{0\}$  and  $d\hat{\mu}$  is a singular bounded nonnegative measure such that  $\text{supp}(d\hat{\mu}) \subset \text{null}(\hat{Q})$ .

*Proof.* Following the same path as in the proof of Theorem 4.5, we now consider the Lagrangian

$$\begin{aligned} L(d\mu, q) &= \mathbb{I}_p(d\mu) + \sum_{k=1}^n q_k \left( c_k - \int_K \alpha_k (\Phi dx + d\hat{\mu}) \right) \\ &= \int_K P(x) \log \Phi(x) dx + \langle c, q \rangle - \int_K Q \Phi dx - \int_K Q d\hat{\mu}. \end{aligned}$$

Note that for  $(d\mu, q)$  to be a saddle point it is necessary that  $q \in \bar{\mathfrak{P}}_+$  and  $\text{supp}(d\hat{\mu}) \subset \text{null}(\hat{Q})$ . This means that the last term must disappear for any saddle point candidate, and we are therefore left with a function that is identical to the Lagrangian in the proof of Theorem 4.5. The proof then follows along the same lines as that of Theorem 4.5. Also note that the existence of  $d\hat{\mu}$  is ensured by Theorem 5.1.  $\square$

It is interesting to note that the functional (5.4) is concave but not strictly concave, since the value does not depend in the singular part  $d\hat{\mu}$ . Therefore it is not surprising that the optimal singular measure  $d\hat{\mu}$  is not guaranteed to be unique. Moreover the function (5.4) can in fact be seen as a Kullback-Leibler-like divergence index between the two measures  $dp(x) := P(x)dx$  and  $d\mu(x)$  [18]. In fact, maximizing (5.4) is equivalent to minimizing  $\int_K P \log(P/\Phi) dx$ , and, since  $dp$  is absolutely continuous with respect to  $d\mu$ , we have the Kullback-Leibler-like divergence

$$\mathbb{S}_{KL}(dp||d\mu) := \int_K \log\left(\frac{dp}{d\mu}\right) dp = \int_K P \log\left(\frac{P}{\Phi}\right) dx,$$

where  $(dp/d\mu) = P/\Phi$  is the Radon-Nikodym derivative [37, p. 553-554].

For later reference we collect the KKT conditions of Theorem 5.1 in the following corollary.

**Corollary 5.3.** *Let  $c \in \mathfrak{C}_+$ . Then  $\hat{q}$  is the optimal solution to*

$$\min_{q \in \bar{\mathfrak{P}}_+} \mathbb{J}_p^c(q) \tag{5.5}$$

*if and only if*

$$\hat{q} \in \bar{\mathfrak{P}}_+, \quad \hat{c} \in \partial \mathfrak{C}_+ \tag{5.6a}$$

$$c_k = \int_K \alpha_k \frac{P}{\hat{Q}} dx + \hat{c}_k, \quad k = 0, 1, \dots, n \tag{5.6b}$$

$$\langle \hat{c}, \hat{q} \rangle = 0 \tag{5.6c}$$

Note that  $\hat{c} = 0$  whenever  $\hat{q} \in \bar{\mathfrak{P}}_+$ , since then  $\hat{Q}(x) > 0$  on all of  $K$ . This is the situation of Corollary 3.5. If  $\hat{q} \in \partial \bar{\mathfrak{P}}_+$ , we have  $\text{supp}(d\hat{\mu}) \subset \text{null}(\hat{Q})$ . Although  $\hat{c}$  in (5.1b) is unique, the measure  $d\hat{\mu}$  may not be unique in general, as can be seen from the following examples.

*Example 5.4.* Next we provide a one-dimensional example where  $d\hat{\mu}$  is not unique. Let  $K = [0, 1]$ ,  $\alpha_1 = 1$  and  $\alpha_2(x) = (1-x)(\cos(\frac{x}{1-x}) + 1)$ . First note that  $\alpha_2(1) := \lim_{x \rightarrow 1} \alpha_2(x) = 0$ , and thus  $\alpha_2$  is continuous on all of  $K$ . Since  $\alpha_2(x) \in [0, 2]$  for all  $x \in K$ , we have  $\bar{\mathfrak{P}}_+ = \{q \in \mathbb{R}^2 \mid q_1 \geq 0, q_2 \geq -\frac{1}{2}q_1\}$ . Now take  $c = (2, \gamma)'$ , where  $\gamma := \int_K \alpha_2 dx$ . Since  $\alpha_2 \geq 0$  but not identically zero,  $\gamma > 0$ . Moreover,

$$\gamma < \int_0^1 \cos\left(\frac{x}{1-x}\right) dx + 1 < 2.$$

Hence  $0 < \gamma < 2$ . Let  $q \in \bar{\mathfrak{P}}_+ \setminus \{0\}$  be arbitrary. If  $q_1 = 0$ , we have  $q_2 > 0$ , and hence  $\langle c, q \rangle = \gamma q_2 > 0$ . If  $q_1 > 0$ , we have  $\langle c, q \rangle = 2q_1 + \gamma q_2 \geq (2 - \frac{\gamma}{2})q_1 > 0$ .

Therefore  $\langle c, q \rangle > 0$  for all  $q \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ , and hence  $c \in \mathfrak{C}_+$ . Next taking  $P = \hat{Q} = \alpha_2$ , we have  $P(1) = \hat{Q}(1) = 0$ , i.e.,  $p = \hat{q} = (0, 1)' \in \partial\mathfrak{P}_+$ . Then forming

$$\begin{aligned}\hat{c}_1 &= c_1 - \int_0^1 \alpha_1 \frac{P}{\hat{Q}} dx = 2 - 1 = 1 \\ \hat{c}_2 &= c_2 - \int_0^1 \alpha_2 \frac{P}{\hat{Q}} dx = \gamma - \gamma = 0\end{aligned}$$

$\langle \hat{c}, q \rangle = q_1 \geq 0$  for all  $q \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ , and hence  $\hat{c} \in \bar{\mathfrak{C}}_+$ . Moreover,  $\langle \hat{c}, \hat{q} \rangle = \hat{q}_1 = 0$ , so  $\hat{c} \in \partial\mathfrak{C}_+$ . Consequently, by Corollary 5.3,  $\hat{q}$  is the minimizer of  $\mathbb{J}_p^c$ , and hence the support of the measure  $d\hat{\mu}$  in (5.1b) is contained in  $\text{null}(\hat{Q}) = \text{null}(\alpha_2)$ . However  $\alpha_2(t)$  have infinitely many zeros, and any measure  $d\mu = dx + d\hat{\mu}$  such that  $\hat{\mu}(K) = 1$  and  $\text{supp}(d\hat{\mu}) \subset \text{null}(\hat{Q})$  is a solution.

*Example 5.5.* It is easy to construct a small two-dimensional example for which the measure  $d\hat{\mu}$ , representing  $\hat{c}$ , is not unique. Let  $K = [0, 1]^2$  and  $\alpha(x) = (1, x_1, x_2, x_1 x_2)'$ , and define

$$c = \int_K \alpha \frac{P}{\hat{Q}} dx + \hat{c},$$

where  $P(x) = x_1$ ,  $\hat{Q} = x_1(1 + x_2)$ , and  $\hat{c} = \int_K \alpha \delta(x_1) dx = (1, 0, 1/2, 0)'$ . Clearly  $c \in \mathfrak{C}_+$  and  $\hat{q} \in \bar{\mathfrak{P}}_+$ . Moreover,  $\hat{c} \in \partial\mathfrak{C}_+$ . In fact,

$$\langle \hat{c}, \hat{q} \rangle = \int_K \hat{Q} d\hat{\mu} = \int_0^1 x_1 \delta(x_1) dx_1 \int_0^1 (1 + x_2) dx_2 = 0.$$

Therefore  $(\hat{c}, \hat{q})$  satisfies the KKT conditions (5.6) and is the unique minimizer of  $\mathbb{J}_p^c$  (Corollary 5.3). However,  $\text{null}(\hat{Q})$  is the whole line  $x_1 = 0$ , and any measure  $d\hat{\mu}$  with mass 1 and support constrained to  $x_1 = 0$  such that  $\int_K x_2 d\hat{\mu} = 1/2$  is a solution. Hence there are infinitely many ways to select  $d\hat{\mu}$ .

**Condition 5.6.** *The vectors  $\alpha(x_1), \alpha(x_2), \dots, \alpha(x_m)$  are linearly independent, where  $x_1, x_2, \dots, x_m$  are the points where the optimal polynomial (5.1) have zeros, i.e.,  $\hat{Q}(x_j) = 0$ .*

**Proposition 5.7.** *Suppose that Condition 5.6 holds. Then the measure  $d\hat{\mu}$  in Theorem 5.1 is unique. Moreover,*

$$d\hat{\mu} = \sum_{j=1}^m a_j \delta(x - x_j) dx \tag{5.7}$$

for some  $a_1, \dots, a_m \in \mathbb{R}^n$ , where  $m \leq n$ .

*Proof.* Inserting (5.7) into (5.1b) yields

$$\hat{c}_k = \sum_{j=1}^m \alpha_k(x_j) a_j,$$



which has a unique solution  $(a_1, a_2, \dots, a_m)$  if  $\alpha(x_1), \alpha(x_2), \dots, \alpha(x_m)$  are linearly independent.  $\square$

*Remark 5.8.* In the one-dimensional case ( $d = 1$ ),  $\alpha(x_1), \alpha(x_2), \dots, \alpha(x_m)$  are linearly independent for all distinct points  $x_1, x_2, \dots, x_m$  such that  $m \leq n - 1$  in any T-system [23, 26], for example, the trigonometric and power moment problems, and also the Herglotz basis used in Nevanlinna-Pick interpolation. In these cases the zero set of  $Q$  will always satisfy the linear independence property of Proposition 5.7, resulting in a unique  $d\hat{\mu}$ .

**Theorem 5.9.** *Let  $(c, p) \in \mathfrak{C}_+ \times (\tilde{\mathfrak{P}}_+ \setminus \{0\})$ . Suppose that Condition 5.6 holds for the minimizer  $\hat{Q}$  of (5.5) and let  $d\mu = P/\hat{Q}dx + d\hat{\mu}$  be the unique corresponding measure in (5.6). Moreover, let  $(p_k)$  be a sequence in  $\tilde{\mathfrak{P}}_+$  such that  $p_k \rightarrow p$  as  $k \rightarrow \infty$ , and let  $d\mu_k = (P_k/\hat{Q}_k)dx + d\hat{\mu}_k$  be measure (5.6) corresponding to the minimizer of  $\mathbb{J}_{p_k}^c$ . Then  $d\mu_k \rightarrow d\mu$  in weak\* as  $k \rightarrow \infty$ .*

*Proof.* We want to show that

$$\int_K f d\mu_k \rightarrow \int_K f d\mu$$

for an arbitrary  $f \in C(K)$ . We may choose  $\rho(x) = r'\alpha(x)$  so that  $\rho(x) = f(x)$  for  $x \in \text{null}(\hat{Q})$ . In fact, since  $\alpha(x_1), \alpha(x_2), \dots, \alpha(x_m)$  are linearly independent (Condition 5.6), it follows that the system of linear equations

$$r' [\alpha(x_1) \quad \alpha(x_2) \quad \cdots \quad \alpha(x_m)] = [f(x_1) \quad f(x_2) \quad \cdots \quad f(x_m)]$$

has a solution. Then, setting  $g = f - \rho$ ,  $\text{null}(\hat{Q}) \subset \text{null}(g)$ . Moreover, since  $d\mu_k$  and  $d\mu$  both satisfy the moment condition (2.1), we have  $\int_K \rho d\mu_k = \int_K \rho d\mu = \langle c, r \rangle$ , and hence it is sufficient to show that

$$\int_K g d\mu_k \rightarrow \int_K g d\mu. \quad (5.8)$$

Next, fix  $\epsilon > 0$ , and choose  $M$  so that  $\sup_k \mu_k(K) \leq M$  and  $\mu(K) \leq M$ . Let  $B_\delta := \{x_0 + x_1 \in K \mid x_0 \in \text{null}(\hat{Q}), \|x_1\|_2 < \delta\}$ , where  $\delta > 0$  is chosen so that  $|g(x)| < \epsilon/(2M)$  on  $B_\delta$ , which can be done since  $g$  is continuous and  $g(x) = 0$  on  $\text{null}(\hat{Q})$ . By Theorem 5.1,  $\text{supp}(d\hat{\mu}) \subset B_\delta$ . Also, since  $\hat{Q}_k \rightarrow \hat{Q}$  uniformly (Theorem 4.4), it follows that, for  $k$  sufficiently large,  $\text{null}(\hat{Q}_k) \subset B_\delta$ , and hence,  $\text{supp}(d\hat{\mu}_k) \subset B_\delta$  (Theorem 5.1). Thus for  $k$  large enough, we have

$$\begin{aligned} \left| \int_K g(d\mu_k - d\mu) \right| &= \left| \int_{B_\delta} g(d\mu_k - d\mu) + \int_{K \setminus B_\delta} g(P_k/\hat{Q}_k - P/\hat{Q})dx \right| \\ &\leq 2M \max_{x \in B_\delta} |g(x)| + \int_{K \setminus B_\delta} |g|dx \max_{x \in K \setminus B_\delta} |P_k/\hat{Q}_k - P/\hat{Q}|. \end{aligned} \quad (5.9)$$

The first term is bounded by  $\epsilon$  (by the definition of  $B_\delta$ ) and the second term tends to zero since  $P_k/\hat{Q}_k \rightarrow P/\hat{Q}$  uniformly on  $K \setminus B_\delta$ . Since  $\epsilon > 0$  is arbitrary, the limit of  $|\int_K g(d\mu_k - d\mu)|$  as  $k \rightarrow \infty$  of (5.9) is zero, and weak\* convergence  $d\mu_k \rightarrow d\mu$  follows.  $\square$

In many classical one-dimensional moment problems, such as the power moment problem and the trigonometric moment problem, there will be cancellation of common factors in  $P$  and  $Q$ . With more general basis functions this is not necessarily the case. In multidimensional generalizations of the classical problems such cancellation may or may not occur. Example 5.5 shows a situation where there is cancellation, whereas no cancellation occurs in the next simple example.

*Example 5.10.* Consider a two-dimensional power moment problem on  $K = [0, 1]^2$  with basis functions  $\alpha(x) = (1, x_1, x_1^2, x_2, x_2^2, x_1x_2)'$ . The polynomials  $P(x) = x_1^2 + x_2^2$  and  $Q(x) = x_1 + x_1^2 + x_2 + x_2^2$  are irreducible with a common zero in  $x = (0, 0)'$ . Moreover,

$$\lim_{x \rightarrow 0} \frac{P(x)}{Q(x)} = 0,$$

showing that it is integrable. Set  $c := \int_K \alpha \frac{P}{Q} dx$ . Then  $\langle c, r \rangle = \int_K R \frac{P}{Q} dx > 0$  for all  $r \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$ , so  $c \in \mathfrak{C}_+$ . Consequently, we have an example where  $\hat{c} = 0$  and both  $p$  and  $q$  belong to  $\partial\mathfrak{P}_+$ , but there is no cancellation.

## 6. Moments on the boundary

If  $c \in \partial\mathfrak{C}_+$ , there is a  $q_0 \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$  such that  $\langle c, q_0 \rangle = 0$ . Then  $\mathbb{J}_p^c(\lambda q_0) \rightarrow -\infty$  as  $\lambda \rightarrow \infty$ . Consequently, the functional (4.1) has no minimum. Then  $d\mu$  in (1.1) cannot have a rational part. In fact, if  $d\mu$  is given by (1.2), then

$$\langle c, q_0 \rangle = \int_K Q_0 \frac{P}{Q} dx + \int_K Q_0 d\hat{\mu} = 0. \quad (6.1)$$

Since the first term is positive and the second in nonnegative, there could be no rational part in  $d\mu$ . Moreover,  $d\hat{\mu}$  must have support in  $\text{null}(Q_0)$ . More precisely, we have the following representation (c.f., Appendix A [31]).

**Proposition 6.1.** *Suppose that  $\mathfrak{P}_+ \neq \emptyset$ . Then, for any  $c \in \partial\mathfrak{C}_+$ , there exists a  $d\mu \in \mathfrak{M}_c$  with support in at most  $n - 1$  points in  $K$ .*

*Proof.* Since  $c \in \bar{\mathfrak{C}}_+$ , it follows from Proposition 2.4 that there is a  $d\mu \in \mathfrak{M}_c$  with support in at most  $n$  points  $x_1, x_2, \dots, x_n$ . Then  $d\mu = \sum_{\ell=1}^n \beta_\ell d\mu_\ell$  for some nonnegative coefficients  $\beta_\ell$ , where  $d\mu_\ell = \delta(x - x_\ell)$  is the Dirac measure, for  $\ell = 1, 2, \dots, n$ . If  $\beta_\ell = 0$  for some  $\ell$ , then  $d\mu$  has support in at most  $n - 1$  points, so only the case that  $\beta_\ell > 0$  for  $\ell = 1, 2, \dots, n$ , remains. Now, since  $c \in \partial\mathfrak{C}_+$ , there is a  $p \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$  such that  $\langle c, p \rangle = 0$ , i.e.,

$$0 = \langle c, p \rangle = p' \int_K \alpha \sum_{\ell=1}^n \beta_\ell d\mu_\ell = p' \sum_{\ell=1}^n \alpha(x_\ell) \beta_\ell$$

and hence  $\{\alpha(x_1), \dots, \alpha(x_n)\}$  are linearly dependent. Therefore one of the measures  $d\mu_\ell$  can be eliminated in the representation

$$c = \int_K \alpha \sum_{\ell=1}^n \beta_\ell d\mu_\ell = \sum_{\ell=1}^n \alpha(x_\ell) \beta_\ell,$$

proving that only a measure  $d\mu$  with support in  $n - 1$  is needed.  $\square$

Next we provide the deferred proof of Theorem 2.3, which is based on Theorem 5.1 only.

*Proof of Theorem 2.3.* From Theorem 5.1 we know that a solution of (2.1) exists for all  $c \in \mathfrak{C}_+$ . Hence it just remains to show that there is a solution for all  $c \in \partial\mathfrak{C}_+$ . For any  $c \in \partial\mathfrak{C}_+$ , let  $c_\varepsilon = c + \varepsilon c_0$ , where  $c_0 \in \mathfrak{C}_+$ . The existence of such a  $c_0$  is insured by Proposition 2.1 since  $\mathfrak{P}_+ \neq \emptyset$ . Then  $c_\varepsilon \in \mathfrak{C}_+$  for all  $\varepsilon > 0$ . Therefore, by Theorem 5.1 and the fact that  $\mathfrak{P}_+ \neq \emptyset$ , we know that there exist a bounded measure  $d\mu_\varepsilon \in C(K)^*$  of the form  $d\mu_\varepsilon = (P/Q_\varepsilon)dx + d\hat{\mu}_\varepsilon$  such that

$$c_\varepsilon = \int_K \alpha d\mu_\varepsilon$$

for some  $p \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ . Now, since  $\mathfrak{P}_+ \neq \emptyset$ , there is a  $p_0 \in \mathfrak{P}_+$ , and there is a  $\delta > 0$  such that  $P_0(x) \geq \delta$  for all  $x \in K$ . Then

$$\langle c_\varepsilon, p_0 \rangle = \int_K P_0 d\mu_\varepsilon \geq \mu_\varepsilon(K)\delta,$$

However,  $\langle c_\varepsilon, p_0 \rangle = \langle c, p_0 \rangle + \varepsilon \langle c_0, p_0 \rangle \leq M$  for some  $M$  and sufficiently small  $\varepsilon$ . Hence  $\mu_\varepsilon(K) \leq M/\delta$  for all such  $\varepsilon$ . Therefore there is a subsequence  $\{d\mu_{\varepsilon_j}\}$  that converges to some  $d\mu$  in weak\* [30, p. 128], [38, p. 246], and consequently

$$c = \lim_{\varepsilon \rightarrow 0} c_\varepsilon = \lim_{j \rightarrow \infty} \int_K \alpha d\mu_{\varepsilon_j} = \int_K \alpha d\mu$$

which proves Theorem 2.3.  $\square$

## 7. Appendix

The following lemma, showing that  $\mathbb{J}_p^c$  is bounded from below, is a direct generalization of Proposition 2.1 in [6], and the proof follows along the same lines as in [6].

**Lemma 7.1.** *Let  $(c, p) \in \mathfrak{C}_+ \times \bar{\mathfrak{P}}_+ \setminus \{0\}$ . Then there are constants  $\epsilon_c, \epsilon_p > 0$  such that*

$$\mathbb{J}_p^c(q) \geq \epsilon_c \|Q\|_\infty - \epsilon_p \log \|Q\|_\infty \quad (7.1)$$

for all  $q \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ .

*Proof.* The linear form  $\langle c, q \rangle$  has a minimum,  $m_c$ , on the compact set  $\{q \in \bar{\mathfrak{P}}_+ \mid \|q\|_\infty = 1\}$ . Since  $c \in \mathfrak{C}_+$ ,  $m_c > 0$ . Hence for an arbitrary  $q \in \bar{\mathfrak{P}}_+ \setminus \{0\}$  we have

$$\langle c, q \rangle = \langle c, \frac{q}{\|q\|_\infty} \rangle \|q\|_\infty \geq m_c \|q\|_\infty.$$

Next we observe that

$$\|Q\|_\infty = \max_{x \in K} Q = \max_{x \in K} \left\{ \sum_{k=0}^n q_k \alpha_k \right\} \leq \sum_{k=0}^n |q_k| |\alpha_k| \leq M \|q\|_\infty,$$

where  $M := \max_{k,x} |\alpha_k|$ . This maximum exists and is positive and finite, since the basis functions  $\alpha_1, \alpha_2, \dots, \alpha_n$  are continuous and  $K$  is compact. Consequently, taking  $\epsilon_c := m_c/M > 0$ , we obtain

$$\langle c, q \rangle \geq \epsilon_c \|Q\|_\infty.$$

Next we consider the integral part of  $\mathbb{J}_p^c$ , namely

$$\int_K P \log Q dx = \int_K P \log \left( \frac{Q}{\|Q\|_\infty} \right) dx + \log \|Q\|_\infty \int_K P dx.$$

Since  $p \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$  is fixed,  $\epsilon_p := \int_K P dx > 0$ . Therefore

$$\mathbb{J}_p^c(q) \geq \epsilon_c \|Q\|_\infty - \int_K P \log \left( \frac{Q}{\|Q\|_\infty} \right) dx - \epsilon_p \log (\|Q\|_\infty).$$

However, the integrand in the second term is nonpositive, and hence (7.1) follows.  $\square$

**Lemma 7.2.** *Let  $c \in \mathfrak{C}_+$ , and let  $\mathbb{J}_p^c(q) : \tilde{\mathfrak{P}}_+ \rightarrow \mathbb{R} \cup \{\infty\}$  be the functional (4.1). Then the optimal value  $\mathbb{J}_p^c(\hat{q}) = \min_{q \in \tilde{\mathfrak{P}}_+} \mathbb{J}_p^c(q)$  is continuous in  $p$  over  $\tilde{\mathfrak{P}}_+ \setminus \{0\}$ .*

*Proof.* Let  $p_1, p_2 \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$  be arbitrary, and let  $q_1, q_2 \in \tilde{\mathfrak{P}}_+$  be the unique minimizers of  $\mathbb{J}_{p_1}^c$  and  $\mathbb{J}_{p_2}^c$ , respectively. Choose a  $q_0 \in \tilde{\mathfrak{P}}_+$  such that  $0 < \|Q_0\|_\infty < \infty$ . Then, by optimality,

$$\mathbb{J}_{p_1}^c(q_1) \leq \mathbb{J}_{p_1}^c(q_2 + \varepsilon q_0) \quad (7.2a)$$

$$\mathbb{J}_{p_2}^c(q_2) \leq \mathbb{J}_{p_2}^c(q_1 + \varepsilon q_0) \quad (7.2b)$$

for all  $\varepsilon > 0$ . Therefore, if we could show that, for any  $\delta > 0$ ,

$$|\mathbb{J}_{p_2}^c(q_1 + \varepsilon q_0) - \mathbb{J}_{p_1}^c(q_1)| \leq \delta \quad (7.3a)$$

$$|\mathbb{J}_{p_1}^c(q_2 + \varepsilon q_0) - \mathbb{J}_{p_2}^c(q_2)| \leq \delta \quad (7.3b)$$

for  $\|p_2 - p_1\|$  sufficiently small, we would have

$$\mathbb{J}_{p_2}^c(q_2) - \delta \leq \mathbb{J}_{p_1}^c(q_1) \leq \mathbb{J}_{p_2}^c(q_2) + \delta,$$

and the lemma would follow. To prove this, form

$$\begin{aligned} & |\mathbb{J}_{p_2}^c(q_1 + \varepsilon q_0) - \mathbb{J}_{p_1}^c(q_1)| \\ &= \left| \langle c, \varepsilon q_0 \rangle - \int_K P_2 \log(Q_1 + \varepsilon Q_0) dx + \int_K P_1 \log Q_1 dx \right| \\ &= \left| \langle c, \varepsilon q_0 \rangle - \int_K P_1 \log \left( 1 + \frac{\varepsilon Q_0}{Q_1} \right) dx - \int_K (P_2 - P_1) \log(Q_1 + \varepsilon Q_0) dx \right| \\ &\leq \varepsilon \left( \langle c, q_0 \rangle + \int_K P_1 \frac{Q_0}{Q_1} dx \right) + \|P_2 - P_1\|_1 \|\log(Q_1 + \varepsilon Q_0)\|_\infty, \end{aligned}$$

where, by optimality of  $q_1$  with respect to  $\mathbb{J}_{p_1}^c$ ,  $\int_K P_1 \frac{Q_0}{Q_1} dx = \langle c, q_0 \rangle$  is finite. Here we can make the first term less or equal to  $\delta/2$  by choosing  $\varepsilon$  sufficiently small. Then, set  $K(\varepsilon) := 2\|\log(Q_1 + \varepsilon Q_0)\|_\infty$ , and take  $\|P_2 - P_1\|_1 \leq \delta/K(\varepsilon)$ ,

from which (7.3a) follows. The inequality (7.3b) follows by the same line of argument.  $\square$

**Proposition 7.3.** *Let  $\mathbb{J}_p^c(q) : \tilde{\mathfrak{P}}_+ \rightarrow \mathbb{R} \cup \{\infty\}$  be the functional (4.1). Then the function*

$$\hat{q} = \arg \min_{q \in \tilde{\mathfrak{P}}_+} \mathbb{J}_p^c(q)$$

*that maps  $p \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$  to  $\hat{q} \in \tilde{\mathfrak{P}}_+$  is continuous for all values of  $c \in \mathfrak{C}_+$ .*

*Proof.* Let  $(p_k)$  be a sequence in  $\tilde{\mathfrak{P}}_+$  converging to  $p \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$  as  $k \rightarrow \infty$ . Moreover, let  $q_k = \arg \min_{q \in \tilde{\mathfrak{P}}_+} \mathbb{J}_{p_k}^c(q)$  and  $\hat{q} = \arg \min_{q \in \tilde{\mathfrak{P}}_+} \mathbb{J}_p^c(q)$ . Then, by Lemma 4.3, the sequence  $(q_k)$  is bounded, and hence there is a subsequence, which we also call  $(q_k)$  converging to a limit  $q_\infty$ . Assume that  $q_\infty \neq \hat{q}$ , and choose a  $q_0 \in \tilde{\mathfrak{P}}_+$  such that  $0 < \|Q_0\|_\infty < \infty$ . Then

$$\begin{aligned} \mathbb{J}_{p_k}^c(q_k) &= \mathbb{J}_{p_k}^c(q_k + \varepsilon q_0) - \langle c, \varepsilon q_0 \rangle + \int_K P_k \log \left( \frac{Q_k + \varepsilon Q_0}{Q_k} \right) dx \\ &\geq \mathbb{J}_{p_k}^c(q_k + \varepsilon q_0) - \langle c, \varepsilon q_0 \rangle. \end{aligned}$$

Therefore, by Lemma 7.2, we have

$$\mathbb{J}_p^c(\hat{q}) = \lim_{k \rightarrow \infty} \mathbb{J}_{p_k}^c(q_k) \geq \lim_{k \rightarrow \infty} \mathbb{J}_{p_k}^c(q_k + \varepsilon q_0) - \varepsilon \langle c, q_0 \rangle.$$

However  $q_k + \varepsilon q_0 \in \tilde{\mathfrak{P}}_+$ , and, since  $(p, q) \mapsto \mathbb{J}_p^c(p, q)$  is continuous in  $\tilde{\mathfrak{P}}_+$ , we obtain

$$\mathbb{J}_p^c(\hat{q}) \geq \lim_{k \rightarrow \infty} \mathbb{J}_{p_k}^c(q_k + \varepsilon q_0) - \varepsilon \langle c, q_0 \rangle = \mathbb{J}_p^c(q_\infty + \varepsilon q_0) - \varepsilon \langle c, q_0 \rangle \quad (7.4)$$

Since the minimum  $\min_{q \in \tilde{\mathfrak{P}}_+} \mathbb{J}_p^c(q)$  is unique (Theorem 4.4), there is a  $\delta > 0$  such that  $\mathbb{J}_p^c(q_\infty) - \delta > \mathbb{J}_p^c(\hat{q})$ , which together with (7.4) yields

$$\mathbb{J}_p^c(q_\infty) - \delta > \mathbb{J}_p^c(q_\infty + \varepsilon q_0) - \varepsilon \langle c, q_0 \rangle$$

However by letting  $\varepsilon \rightarrow 0$  we get  $-\delta > 0$ , which is a contradiction. Therefore  $\lim_{k \rightarrow \infty} q_k = \hat{q}$ , as claimed.  $\square$

**Proposition 7.4.** *Let  $\hat{q} := \arg \min_{q \in \tilde{\mathfrak{P}}_+} \mathbb{J}_p^c(q)$ , where  $\mathbb{J}_p^c(q) : \tilde{\mathfrak{P}}_+ \rightarrow \mathbb{R} \cup \{\infty\}$  is the functional (4.1). Then the function sending  $c \in \mathfrak{C}_+$  to  $\hat{q}$  is continuous for all values of  $p \in \tilde{\mathfrak{P}}_+ \setminus \{0\}$ .*

*Proof.* The proof is analogous to that of Proposition 7.3. We first prove that the optimal value  $\mathbb{J}_p^c(\hat{q}) = \min_{q \in \tilde{\mathfrak{P}}_+} \mathbb{J}_p^c(q)$  is continuous in  $c$  over  $\mathfrak{C}_+$ . To this end, let  $c_1, c_2 \in \mathfrak{C}_+$  be arbitrary, and let  $q_1, q_2 \in \tilde{\mathfrak{P}}_+$  be the unique minimizers of  $\mathbb{J}_p^{c_1}$  and  $\mathbb{J}_p^{c_2}$ , respectively. Next, exchanging  $\mathbb{J}_{p_1}^c$  for  $\mathbb{J}_{p_1}^{c_1}$  and  $\mathbb{J}_{p_2}^c$  for  $\mathbb{J}_{p_2}^{c_2}$  everywhere in the proof of Lemma 7.2, we see that it just remains to show that the absolute value of

$$\mathbb{J}_p^{c_2}(q_1 + \varepsilon q_0) - \mathbb{J}_p^{c_1}(q_1) \leq \|q_1\| \|c_2 - c_1\| + \varepsilon \left[ \langle c_1, q_0 \rangle + \int_K P \frac{Q_0}{Q_1} dx \right]$$

can be made smaller than some preselected  $\delta > 0$  by choosing an appropriately small  $\varepsilon$ . This follows from the same argument as in Lemma 7.2, thus establishing the continuity of the optimal value.

Next, let  $(c_k)$  be a sequence in  $\mathfrak{C}_+$  converging to  $c \in \mathfrak{C}_+$  as  $k \rightarrow \infty$ , and let  $q_k = \arg \min_{q \in \mathfrak{P}_+} \mathbb{J}_p^{c_k}(q)$  and  $\hat{q} = \arg \min_{q \in \mathfrak{P}_+} \mathbb{J}_p^c(q)$ . Since  $(q_k)$  is bounded, there is a subsequence, also called  $(q_k)$  such that  $q_k \rightarrow q_\infty$ . We assume that  $q_\infty \neq \hat{q}$ , and show that this leads to a contradiction. As in the proof of Proposition 7.3, we first note that

$$\mathbb{J}_p^{c_k}(q_k) \geq \mathbb{J}_p^{c_k}(q_k + \varepsilon q_0) - \varepsilon \langle c_k, q_0 \rangle$$

and hence, by the continuity of the optimal value of  $\mathbb{J}_p^c$  in  $c$ , that

$$\begin{aligned} \mathbb{J}_p^c(\hat{q}) &= \lim_{k \rightarrow \infty} \mathbb{J}_p^{c_k}(q_k) \geq \lim_{k \rightarrow \infty} (\mathbb{J}_p^{c_k}(q_k + \varepsilon q_0) - \varepsilon \langle c_k, q_0 \rangle) \\ &= \mathbb{J}_p^c(q + \varepsilon q_0) - \varepsilon \langle c, q_0 \rangle, \end{aligned}$$

where we have used the fact that  $\mathbb{J}_p^c$  is continuous in  $(c, q) \in \mathfrak{C}_+ \times \mathfrak{P}_+$ . Then letting  $\varepsilon \rightarrow 0$  leads to a contradiction, proving that  $q_\infty = \hat{q}$ , as claimed.  $\square$

## 8. Errata

The proof of the Lemma 3.1 in [6] contains an error emanating from an incorrect use of the dominate convergence theorem, which invalidates Theorem 1.10 in [6]. The correct statement is the one-dimensional version of Theorem 5.1 in the present paper.

## References

- [1] C.I. Byrnes, P. Enqvist, and A. Lindquist. Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach. *SIAM Journal on Control and Optimization*, 41(1):23–59, 2002.
- [2] C.I. Byrnes, T.T. Georgiou, and A. Lindquist. A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. *IEEE Transactions on Automatic Control*, 46(6):822–839, 2001.
- [3] C.I. Byrnes, S.V. Gusev, and A. Lindquist. A convex optimization approach to the rational covariance extension problem. *SIAM Journal on Control and Optimization*, 36:211–229, 1998.
- [4] C.I. Byrnes, S.V. Gusev, and A. Lindquist. From finite covariance windows to modeling filters: A convex optimization approach. *SIAM Review*, 43(4):pp. 645–675, 2001.
- [5] C.I. Byrnes and A. Lindquist. A convex optimization approach to generalized moment problems. In Y. Oishi K. Hashimoto and Y. Yamamoto, editors, *Control and Modeling of Complex Systems: Cybernetics in the 21st Century: Festschrift in Honor of Hidenori Kimura on the Occasion of his 60th Birthday*, pages 3–21. Birkhäuser, 2003.
- [6] C.I. Byrnes and A. Lindquist. The generalized moment problem with complexity constraint. *Integral Equations and Operator Theory*, 56(2):163–180, 2006.
- [7] C.I. Byrnes and A. Lindquist. Important moments in systems and control. *SIAM Journal on Control and Optimization*, 47(5):2458–2469, 2008.

- [8] C.I. Byrnes and A. Lindquist. The moment problem for rational measures: convexity in the spirit of Krein. In *Modern Analysis and Application: Mark Krein Centenary Conference, Vol. I: Operator Theory and Related Topics*, pages 157–169. Birkhäuser, 2009.
- [9] C.I. Byrnes, A. Lindquist, S.V. Gusev, and A.S. Matveev. A complete parameterization of all positive rational extensions of a covariance sequence. *IEEE Transactions on Automatic Control*, 40(11):1841–1857, 1995.
- [10] M.P. Ekstrom. *Digital image processing techniques*. Academic Press, 1984.
- [11] A. Ferrante, M. Pavon, and F. Ramponi. Hellinger versus Kullback-Leibler multivariable spectrum approximation. *IEEE Transactions on Automatic Control*, 53(4):954–967, 2008.
- [12] T.T. Georgiou. *Partial Realization of Covariance Sequences*. PhD thesis, 1983. Center for mathematical systems theory, Univeristy of Florida.
- [13] T.T. Georgiou. Realization of power spectra from partial covariance sequences. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(4):438–449, 1987.
- [14] T.T. Georgiou. A topological approach to Nevanlinna-Pick interpolation. *SIAM Journal on Mathematical Analysis*, 18:1248–1260, 1987.
- [15] T.T. Georgiou. The interpolation problem with a degree constraint. *IEEE Transactions on Automatic Control*, 44(3):631–635, 1999.
- [16] T.T. Georgiou. Solution of the general moment problem via a one-parameter imbedding. *IEEE Transactions on Automatic Control*, 50(6):811–826, 2005.
- [17] T.T. Georgiou. Relative entropy and the multivariable multidimensional moment problem. *IEEE Transactions on Information Theory*, 52(3):1052–1066, 2006.
- [18] T.T. Georgiou and A. Lindquist. Kullback-leibler approximation of spectral density functions. *IEEE Transactions on Information Theory*, 49(11):2910–2917, 2003.
- [19] I.V. Girsanov. *Lectures on mathematical theory of extremum problems*. Springer, 1972.
- [20] J. Hadamard. Sur les correspondances ponctuelles. *Oeuvres, Editions du Centre Nationale de la Recherche Scientifique, Paris*, pages 383–384, 1968.
- [21] C.V. Jakowatz, D.E. Wahl, P.H. Eichel, D.C. Ghiglia, and P.A. Thompson. *Spotlight-mode synthetic aperture radar: a signal processing approach*, volume 101. Kluwer Academic Publishers Boston, 1996.
- [22] R. Kalman. Realization of covariance sequences. In *Toeplitz memorial conference*, 1981. Tel Aviv, Israel.
- [23] S. Karlin and W.J. Studden. *Tchebycheff systems: With applications in analysis and statistics*, volume 376. Interscience Publishers New York, 1966.
- [24] J. Karlsson and T.T. Georgiou. Signal analysis, moment problems & uncertainty measures. In *44th IEEE Conference on Decision and Control (CDC), and European Control Conference (ECC)*, pages 5710–5715, Dec 2005.
- [25] M.G. Krein and D. Milman. On extreme points of regular convex sets. *Studia Mathematica*, 9(1):133–138, 1940.

- [26] M.G. Krein and A.A. Nudelman. *The Markov moment problem and exextremal problems*, volume Fifty of *Translations of Mathematical monographs*. American Mathematical Society, Providence, R.I., 1977.
- [27] S.W. Lang and J.H. McClellan. Spectral estimation for sensor arrays. In *Proceedings of the First ASSP Workshop on Spectral Estimation*, pages 3.2.1–3.2.7, 1981.
- [28] S.W. Lang and J.H. McClellan. Multidimensional MEM spectral estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(6):880–887, Dec. 1982.
- [29] S.W. Lang and J.H. McClellan. Spectral estimation for sensor arrays. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(2):349–358, Apr. 1983.
- [30] D.G. Luenberger. *Optimization by Vector Space Methods*. Wiley, New York, 1969.
- [31] J.H. McClellan and S.W. Lang. Mult-dimensional MEM spectral estimation. In *Proceedings of the Institute of Acoustics "Spectral Analysis and its Use in Underwater Acoustics": Underwater Acoustics Group Conference, Imperial College, London, 29-30 April 1982*, pages 10.1–10.8, 1982.
- [32] P. Milanfar, W.C. Karl, and A.S. Willsky. A moment-based variational approach to tomographic reconstruction. *IEEE Transactions on Image Processing*, 5(3):459–470, 1996.
- [33] H.I. Nurdin. New results on the rational covariance extension problem with degree constraint. *Systems & Control Letters*, 55(7):530 – 537, 2006.
- [34] H.I. Nurdin and A. Bagchi. On the solution of the rational covariance extension problem corresponding to pseudo-polynomials having boundry zeros. *IEEE Transactions on Automatic Control*, 51(2):350–355, 2006.
- [35] M. Pavon and A. Ferrante. On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities. *IEEE Transactions on Automatic Control*, 51(4):639–644, 2006.
- [36] M. Pavon and A. Ferrante. On the geometry of maximum entropy problems. *SIAM Review*, 55(3):415–439, 2013.
- [37] A. Rényi. On measures of entropy and information. In *Fourth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 547–561, 1961.
- [38] W. Rudin. *Real and complex analysis*. McGraw-Hill, New York, 1987.
- [39] M. Spivak. *Calculus on Manifolds*. Benjamin Cummings, Menlo Park, California, 1965.

Johan Karlsson  
 Department of Mathematics  
 Royal Institute of Technology  
 SE-100 44 Stockholm  
 Sweden  
 e-mail: [johan.karlsson@math.kth.se](mailto:johan.karlsson@math.kth.se)



Anders Lindquist  
Departments of Automation and Mathematics  
Shanghai Jiao Tong University  
200240 Shanghai  
China  
and  
Department of Mathematics  
Royal Institute of Technology  
SE-100 44 Stockholm  
Sweden  
e-mail: [alq@kth.se](mailto:alq@kth.se)

Axel Ringh  
Department of Mathematics  
Royal Institute of Technology  
SE-100 44 Stockholm  
Sweden  
e-mail: [aringh@kth.se](mailto:aringh@kth.se)